

# Horses for Courses: Mean-Variance for Asset Allocation and $1/N$ for Stock Selection

Emmanouil Platanakis<sup>#</sup>, Charles Sutcliffe<sup>\*</sup> and Xiaoxia Ye<sup>†</sup>

13 May 2019

- # School of Management, University of Bath, Claverton Down, Bath, BA2 7AY, UK, email: [E.Platanakis@bath.ac.uk](mailto:E.Platanakis@bath.ac.uk) Tel: +44(0) 1274 235311, ORCID: 0000-0003-1493-6339
- \* The ICMA Centre, Henley Business School, University of Reading, PO Box 242, Reading RG6 6BA, UK, email: [c.m.s.sutcliffe@rdg.ac.uk](mailto:c.m.s.sutcliffe@rdg.ac.uk) Tel +44(0) 118 378 6117, ORCID 0000-0003-0187-487X
- † Management School, University of Liverpool, Chatham Street, Liverpool L69 7ZH, UK, email: [Xiaoxia.Ye@liverpool.ac.uk](mailto:Xiaoxia.Ye@liverpool.ac.uk) Tel +44 (0)151 795 7638

Keywords: asset allocation; stock selection; mean-variance; naive diversification; portfolio theory

JEL - G11

We wish to thank Chris Brooks (Reading) and Robert Jarrow (Cornell) for their comments on an earlier draft.

11,697 words

## **Horses for Courses: Mean-Variance for Asset Allocation and $1/N$ for Stock Selection**

### **Abstract**

For various organizational reasons, large investors typically split their portfolio decision into two stages - asset allocation and stock selection. We hypothesise that mean-variance models are superior to equal weighting for asset allocation, while the reverse applies for stock selection, as estimation errors are less of a problem for mean-variance models when used for asset allocation than for stock selection. We confirm this hypothesis in separate analyses of US and international equities using four different types of mean-variance model (Bayes-Stein, Black-Litterman, Bayesian diffuse prior and Markowitz), a range of parameter settings, and a simulation analysis calibrated to US data.

98 words

## Horses for Courses: Mean-Variance for Asset Allocation and $1/N$ for Stock Selection

Markowitz (1952) proposed the formation of portfolios as a single problem across all the available assets. However, in reality, most large investors follow a two-stage process. In the first stage (*asset allocation*) a high-level committee allocates the funds between asset classes, e.g. domestic equities, foreign equities, bonds, properties, commodities, hedge funds, private equity, etc. This often takes the form of setting a benchmark asset allocation against which performance is judged (the reference portfolio). In the second stage (*stock selection*) fund managers specialising in each asset class select individual assets from within their asset class. This two-stage process has become more explicit with the shift of large investors such as pension funds and insurers away from balanced fund managers who both asset allocate and stock select, to dividing their funds between multiple specialist managers confined to stock selection within their allocated asset class<sup>1</sup>. Blake et al. (2013) documented this switch from balanced to multiple specialist managers, with the relative proportion using a balanced fund manager falling from 98% in 1984 to 26% in 2004.

There are a number of reasons why large investors have adopted a two-stage process. Allocating assets across a small number of asset classes is something that can be addressed by a high-level committee. Unlike external fund managers, in banks, insurance companies and pension plans etc. such a committee has responsibility for the investor's asset-liability performance, and can allow for correlations between the investor's liabilities and asset classes when choosing the asset allocation and asset-liability risk exposure. Following Brinson et al (1986) there is a widespread view that asset allocation is much more important than stock selection in determining portfolio performance<sup>2</sup>, which also encourages investors to set the reference portfolio. However, although a high-level committee can set the reference

---

<sup>1</sup> Fiduciary management, where the investor outsources the entire investment process, has increased in popularity in recent years, and this also follows a two-stage process.

<sup>2</sup> This view has been shown to depend on the definition of the question (Ibbotson and Kaplan, 2000).

portfolio, it is not practical for them to make a choice between thousands of companies, bonds etc., so they delegate this to specialist managers. The high-level committee probably meets infrequently and the reference portfolio may apply for several years, while stock selection is something that changes on a much more frequent basis, again leading to a need to delegate stock selection<sup>3</sup>. Delegation of stock selection has the advantage that specialist fund managers have greater knowledge and expertise than a high-level committee, and splitting the money among a number of asset managers avoids diseconomies of scale. Employing multiple specialist managers also has the advantages of diversifying across stock selection strategies and manager skills, permitting the choice of the ‘best of breed’ manager within each asset class, and generating competition between asset managers.

There are a number of disadvantages to adopting a two-stage process where stock selection is delegated to multiple specialists. The agency issues involved in such delegation have been investigated by many authors (see Stracca, 2006 for a review), and involve incentivising, rewarding and monitoring the agents (fund managers) to make decisions in accordance with the principal’s preferences. For example, fund managers may be rewarded for exceeding their benchmark, but not penalised for under-performance, leading them to take more risk than desired by the principal. They may also have a much shorter investment horizon than the principal due to their performance being judged over months rather than decades. Splitting portfolio selection into two-stages leads to the loss of some of the benefits of diversification (Sharpe, 1981, Van Binsbergen et al. 2008) as the correlations between individual assets in different asset classes are ignored, and effectively replaced by the average correlation between all the individual assets in each pair of asset classes. In addition, some individual assets fall within the definition of two asset classes. For example, shares in an airport may be bought by an equity manager and an infrastructure manager, and this can lead to over-weighting such assets in the investor’s total portfolio. Finally the management fees and administration costs for multiple

---

<sup>3</sup> While the reference portfolio can be changed at any time, it is probably reviewed by pension plans on the same three year cycle as the actuarial valuation.

specialist managers will usually be larger than for a single balanced manager.

Given the widespread adoption of a two-stage process, investors must think the benefits outweigh the costs. Blake et al (2013) found that, despite the diversification loss, agency problems and higher management fees, this switch has led to better performance due to the avoidance of diseconomies of scale, the creation of competition between fund managers, and the greater knowledge and skill of specialist managers. We argue below there is another powerful reason for this change - improved investment performance due to better portfolio weights.

While portfolio weights can be set without using any formal model, we investigate the performance of different optimization models, e.g. Markowitz, Black-Litterman (BL), Bayes-Stein (BS), Bayesian diffuse prior (BDP) and  $1/N$ , in making the asset allocation and stock selection decisions. This requires the investor and their fund managers to choose a portfolio optimization model, each with its own strengths and weaknesses, when used in a one or two-stage portfolio selection process.

Solving the portfolio problem in a single stage using an optimization technique such as mean-variance (MV) runs into data problems. When a historical estimate of the covariance matrix is used, the length of the time series of asset returns must be greater than the number of assets under consideration to allow estimation of the covariances between possibly thousands of individual assets. Such a time series is unavailable for most assets; and even if a sufficiently long time series is available, it may not be stationary. Therefore only a sub-set of the individual assets can be included in such one-stage optimization models<sup>4</sup>. The problem of insufficient observations to estimate the covariance matrix is much less pressing for a two-stage optimization process. The number of asset classes in the first stage optimization may be in single figures, and the number of individual assets in each asset class at the

---

<sup>4</sup> If the investor is willing to impose conditions on the estimated covariance matrix, a shorter time series can be used. For example, investors can use a single-index model, various types of multi-index model, or a mean correlation model to estimate the correlation matrix for input to a single stage portfolio problem.

second stage will be much smaller than the total number of assets across all asset classes.

Asset classes constitute grouped data, while individual assets do not. If the individual assets within a particular asset class are reasonably homogeneous, the forecast risk, return and correlations for asset classes are likely to be more accurate than those for individual assets (Elton and Gruber, 1971; Lam et al. 1994; Frankfurter and Phillips, 1980). MV and its variants is an error maximiser as it invests primarily in assets with the highest forecast returns, the lowest forecast risk and the lowest estimated correlations (Michaud and Michaud, 2008). There are far fewer asset classes than individual assets within an asset class; and so individual assets offer greater chances of risk and return forecasts that are outliers. This suggests MV is better suited to forming portfolios of assets classes (asset allocation) than portfolios of individual assets (stock selection) as estimation errors and outliers will be less of a problem for asset allocation than for stock selection.

We analyse the relative performance of a version of MV in forming portfolios in the first and second stages, relative to that of naive diversification ( $1/N$ ). We use Bayes-Stein with variance based constraints (VBC) as a representative of the many variations of MV (see section 2). We use this well known model because it is designed to address the well known problem of estimation errors in MV portfolio model inputs<sup>5</sup>. The equal proportions ( $1/N$ ) model is a very simple model that is unaffected by estimation errors, and many recent studies have found that no other model is consistently superior, e.g. DeMiguel et al. (2009a). Therefore we use  $1/N$  as our alternative to MV. Estimation errors are likely to be smaller at the first stage due to grouping individual assets into asset classes, and fewer asset classes than individual assets, leading to a lower chance of outliers. Because estimation errors and outliers are expected to be less of a problem for the first than the second stage, we expect MV-based models to be better than  $1/N$  for asset allocation. We also expect that the estimation errors and associated outliers

---

<sup>5</sup> We also use the Black-Litterman, Bayesian diffuse prior and Markowitz models and upper generalized constraints (UGC) in our robustness checks.

are sufficiently large that the performance of  $1/N$  is better than MV in the second stage when selecting portfolios of individual assets. Therefore our hypothesis is that using MV for asset allocation, and then  $1/N$  for stock selection is superior to the other three possible combinations of these two portfolio models in a two-stage procedure.

We examine this hypothesis in two different ways. First we compare the out-of-sample performance of portfolios of equities (US sector indices and individual equities in our main analysis, and country equity indices and individual equities in our robustness check) formed using a two-stage process employing either MV or  $1/N$  at each stage. We conclude that MV is superior to  $1/N$  for the first stage (asset allocation); and  $1/N$  is preferable to MV for the second stage (stock selection). In other words investors need to use 'different horses for different courses'<sup>6</sup>. We then support these conclusions with a simulation analysis calibrated using US data that allows assets to have non-zero CAPM alphas, a range of robustness checks using US data, and a replication of our core analysis on international data covering seven countries.

DeMiguel et al. (2009a) state that one of the reasons why the  $1/N$  rule performs well on their data sets is that they engage in asset allocation across asset classes, rather than forming portfolios of individual assets. They believe the loss from naive, as opposed to optimal, diversification is smaller when allocating wealth across asset classes due to the lower cross sectional idiosyncratic volatility of asset classes. This conclusion is inconsistent with our finding that, although asset classes have much lower idiosyncratic volatility than individual stocks, MV outperforms  $1/N$ . In DeMiguel et al. (2009a)'s simulation, the returns are simulated using a one-factor model with alpha being strictly zero. In our simulation, following DeMiguel et al. (2009a), we also use a one-factor model and the utility loss function defined by Kan and Zhou (2007) to quantify the out-performance of MV over  $1/N$ . We differ

---

<sup>6</sup> In a different context, Jarrow and Zhao (2006) show that MV is suitable for equity portfolios, while mean-lower partial moment is preferred for fixed income portfolios.

from DeMiguel et al. (2009a) by allowing alpha to be a random variable with zero mean and small variance. In this more realistic setting we find that the conclusions of DeMiguel et al. (2009a) are reversed. The lower is cross sectional idiosyncratic volatility of asset classes, the greater is the superiority of MV over  $1/N$  for portfolios of asset classes, and *vice versa*.

We do not compare the performance of one versus two-stage portfolio formation procedures because the superiority of the two-stage process is affected by factors that cannot be incorporated in an analysis of past asset returns, e.g. the avoidance of diseconomies of scale, the creation of competition between fund managers, the greater knowledge and skill of specialist managers, agency problems, short investment horizons, the definition of asset classes, and management fees and administration costs. There is also the issue that the time series of asset returns may not be long enough to estimate the covariance matrix when MV is applied to all the individual assets.

There is a substantial literature which compares the performance of different portfolio optimization techniques, but this research has not drawn a clear distinction between portfolio optimization models applied to asset classes and individual stocks. A number of studies have applied MV and  $1/N$  to either asset allocation or stock selection problems and obtained results consistent with each of the parts of our hypothesis. For example, Allaj (2017), Durand, et al (2011), Han (2016) and Shigeta (2016) have found that MV has a bigger Sharpe ratio than  $1/N$  for asset classes; while Barroso (2015), Board and Sutcliffe (1994), DeMiguel et al. (2009b), Dickson (2016), Li (2016), Hwang, et al (2018) and Jagannathan and Ma (2003) have found that  $1/N$  has a bigger Sharpe ratio than MV for individual equities. Frankfurter and Phillips (1980) used factor and cluster analysis to form 476 individual US equities into 40 homogeneous groups, or quasi-securities. Within each quasi-security the stock selection used  $1/N$ . They then applied MV to these 40 quasi-securities to select the portfolio of quasi-securities (i.e. asset allocation). They compared the out-of-sample performance of their two-stage approach to that of MV applied directly to the 476 individual assets, i.e. a one-stage model, with the market model used to



estimate the inputs, and conclude that their two-stage approach is superior to a one-stage process. However, they did not compare the performance of different portfolio models in the various stages of a two-stage process.

We are the first to compare the out-of-sample performance of portfolios chosen using different portfolio models for each of the two-stages, and to find that the appropriate portfolio model depends on the nature of the assets being analysed. We are also the first to investigate this question using a simulation which permits assets to have non-zero alphas; which reverses the conclusions of DeMiguel et al (2009a).

Section 2 presents the portfolio models we use in our main analysis, and Section 3 describes our data. Section 4 contains our methodology, and Section 5 has our results. Section 6 sets out the simulation analysis of our hypothesis, and Section 7 has robustness checks on our main results. Finally, Section 8 contains our conclusions.

## **2. Portfolio Models**

### *A. Mean-variance (MV) Model*

The main problem with MV is estimation error. The MV portfolio framework is an optimal strategy in terms of the expected utility if asset returns follow a normal distribution and the parameter estimates are known with certainty, (Hanoch and Levy, 1969). However, the future (true) parameters of the distribution of asset returns are not known with certainty, and their prediction is an extremely difficult task. Historical estimates of returns are often used to compute the optimal asset weights, but this practice usually leads to weak out-of-sample performance with extreme asset weights. This is because marginal deviations from the future (true) parameter estimates have a large effect on the estimated asset weights.

For this reason portfolio optimization strategies, such as Bayes-Stein, have been proposed to decrease the negative effects of estimation errors in the portfolio construction process by changing the way mean returns and the covariance matrix of returns are estimated. Therefore, in the core part of our study we use an enhanced version of the Bayes–Stein shrinkage approach, which was initially proposed by Jorion (1985, 1986). This model is based on the shrinkage estimation of James and Stein (1961), and offers a robust way of dealing with estimation risk. Bayes–Stein has been recognized as a well-established portfolio technique that is robust to estimation risk, and has been widely used in the academic literature, e.g. Board and Sutcliffe (1994), Kan and Zhou (2007), DeMiguel et al. (2009a), DeMiguel et al. (2009b), Tu and Zhou (2011), DeMiguel et al. (2013), Bessler et al. (2017), Platanakis and Sutcliffe (2017), Platanakis et al. (2019), Hwang et al. (2017) and Kan et al. (2016).

Bayes-Stein is based on the idea that estimated returns closer to the norm suffer less from estimation risk, while the estimated returns a long way from the norm are more likely to be due to estimation errors. Hence, Bayes-Stein computes the input parameters as the weighted sum of the sample (historical) returns for each asset in the portfolio and the norm estimate of returns, which is the global minimum variance portfolio (minimum variance portfolio when short-selling is permitted). More specifically, Bayes-Stein computes the column vector of mean returns  $(\boldsymbol{\mu}_{BS})$  as follows:-

$$\boldsymbol{\mu}_{BS} = (1 - g)\boldsymbol{\mu} + g\mu_G \mathbf{1} \quad (1)$$

where the shrinkage factor  $g$ , where  $0 \leq g \leq 1$  is computed by the following equation:-

$$g = \frac{N + 2}{(N + 2) + T(\boldsymbol{\mu} - \mu_G \mathbf{1})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - \mu_G \mathbf{1})}. \quad (2)$$

The parameter  $\mu_G$  denotes the expected return of the minimum variance portfolio when short-selling is allowed, and  $T$  represents the length of the corresponding estimation period. The covariance matrix

of asset returns ( $\Sigma_{BS}$ ) is estimated as follows:-

$$\Sigma_{BS} = \left( \frac{T + \varphi + 1}{T + \varphi} \right) \Sigma + \frac{\varphi}{T(T + \varphi + 1)} \frac{\mathbf{1}\mathbf{1}^T}{\mathbf{1}^T \Sigma^{-1} \mathbf{1}}, \quad (3)$$

where

$$\varphi = \frac{N + 2}{(\boldsymbol{\mu} - \mu_G \mathbf{1})^T \Sigma^{-1} (\boldsymbol{\mu} - \mu_G \mathbf{1})}. \quad (4)$$

We use the Bayes-Stein estimates ( $\boldsymbol{\mu}_{BS}, \Sigma_{BS}$ ), and compute the vector of asset weights (decision variables), denoted by  $\mathbf{x}$ , by maximizing the following quadratic utility function:-

$$U = \mathbf{x}^T \boldsymbol{\mu}_{BS} - \frac{\lambda}{2} \mathbf{x}^T \Sigma_{BS} \mathbf{x} \quad (5)$$

where the parameter  $\lambda$  represents the investor's relative risk aversion. We impose constraints to rule out short-selling ( $\mathbf{x}_i \geq 0, \forall i$ ) as well as constraints for the normalization of portfolio weights

$$\left( \sum_{i=1}^N \mathbf{x}_i = 1 \right).$$

Several academic studies have attempted to decrease the effects of estimation risk by imposing additional constraints on the asset weights to rule out the corner (extreme) solutions produced by estimation errors in the input parameters. For example, Board and Sutcliffe (1994) report that out-of-sample portfolio performance is improved by imposing short-selling constraints, while Jagannathan and Ma (2003) enhance the performance of portfolios by imposing homogenous constraints on the asset weights. DeMiguel et al. (2009a) advocate using portfolio norm constraints, and DeMiguel et al. (2009b) impose upper generalized norm constraints in some of their models. Measured by the Sharpe ratio, Levy and Levy (2014) show that their method of VBCs typically yields a better performance than ten

other portfolio optimization techniques.

For our core analysis we impose the VBCs of Levy and Levy (2014) in an attempt to decrease further the negative effects of estimation errors on the MV input parameters. Since VBCs are anti-proportional to an asset's standard deviation, they constrain more strongly the weights on assets with larger estimation errors as measured by their standard deviation. VBCs are defined as follows:-

$$\left| x_i - \frac{1}{N} \right| \frac{\sigma_i}{\bar{\sigma}} \leq \alpha, \quad \forall i \quad (6)$$

where  $\bar{\sigma}$  is the average standard deviation across all the assets used in the portfolio construction process. The upper bound,  $\alpha$ , is chosen by the investor, and  $\alpha = 0$  is equivalent to the naïve diversification strategy ( $1/N$ ). The effect of the VBCs becomes weaker as the parameter  $\alpha$  increases, and disappears when  $\alpha \rightarrow \infty$ . As a result, our model for optimal diversification in the core analysis is as follows:-

$$\begin{aligned} \max_{\mathbf{x}} \quad & \left\{ \mathbf{x}^T \boldsymbol{\mu}_{BS} - \frac{\lambda}{2} \mathbf{x}^T \boldsymbol{\Sigma}_{BS} \mathbf{x} \right\} \\ \text{s.t.} \quad & \left| x_i - \frac{1}{N} \right| \frac{\sigma_i}{\bar{\sigma}} \leq \alpha, \quad \forall i \\ & x_i \geq 0, \quad \forall i \\ & \sum_{i=1}^N x_i = 1 \end{aligned} \quad (7)$$

#### B. Naïve Diversification ( $1/N$ )

A portfolio weight of  $1/N$  is assigned to each asset, e. g.  $x_i = 1/N$  for all  $i$ . We use  $1/N$  with re-balancing, as in DeMiguel et al. (2009a) rather than  $1/N$  (buy and hold). This method of naïve diversification is very simple, does not require any optimization, is unaffected by estimation errors, and does not depend on the value of the investor's risk aversion.

### 3. Data

Our hypothesis involves allocating money between asset classes, and then selecting assets within each asset class, and the same logic applies to allocating money between industries or counties within a given asset class. In our empirical analysis we consider US stocks, as the data is readily available for investing in companies in different industrial sectors. We treat value-weighted indices of ten US industrial sectors as our asset classes, with asset allocation across the ten industries; and then perform stock selection within each industry. We analyse monthly total returns from January 1994 to August 2017 on ten S&P500 sector indices - information technology, energy, utilities, discretionary consumption, materials, industrials, consumer staples, health care, telecommunications and financials from Datastream. For the second stage we use total returns on the ten companies in each industrial sector index with the largest initial value, except that the telecommunications sector index only comprises three companies<sup>7</sup>. So in total we have 93 companies - see Appendix 1. For risk free returns we use 1-month T-bill returns from Ken French's web site.

### 4. Methodology

We use an out-of-sample procedure, which is more realistic than in-sample tests that assume investors have full knowledge of the distributional parameters of returns. We employ a 12-month expanding estimation window for both the mean return and covariance matrix estimates. Since the length of the estimation window expands over time, an expanding window approach is expected to produce more stable input estimates. We estimate the optimal asset allocation using data available up to and including time  $t$  ( $t \geq 12$ ), and these optimal asset weights are then used to compute the out-of-sample portfolio returns for the next month ( $t+1$ ). This process is repeated by expanding our sample by one-month, re-estimating the portfolio inputs and re-balancing the portfolio until we reach the end of our sample period. Since our data covers the period from January 1994 to August 2017 (284 months), and we use

---

<sup>7</sup> We chose the largest ten companies because institutional investors invest most of their money in large companies, and we need a reasonably small number of assets when estimating the portfolio input parameters.

a 12-month expanding estimation window, we start investing in January 1995 and our out-of-sample period consists of 272 months (January 1995 to August 2017).

We maximize a quadratic utility function using Bayes-Stein estimates for the means and the covariance matrix, in conjunction with non-short selling constraints, normalization of portfolio weights and VBCs, as described in the portfolio optimization problem (eqn. 7). For the core analysis of this study we set the value of the risk aversion parameter ( $\lambda$ ) to 5, which is the value for moderately risk averse investors (e.g. Diris et al. 2015). We set the value of  $\alpha$  involved in the VBCs to 0.15, which is the middle of the range of values used in Levy and Levy (2014). In Section 7 we conduct robustness checks for different levels of risk aversion and different values of  $\alpha$ .

When assessing portfolio performance we allow for transaction costs. Each period we compute the new optimal portfolio of ten asset classes, ten (or three) individual equities or 93 shares, and compare this with the optimal portfolio for the previous period. The total transaction costs are computed using eqn. 8, and subtracted from the out-of-sample gross return on the new portfolio to give the net return.

$$TC_t = \sum_{i=1}^N T_i \left( \left| x_{i,t} - x_{i,t-1}^+ \right| \right), \quad (8)$$

where  $x_{i,t-1}^+$  is the weight (proportion) of the  $i^{\text{th}}$  asset at the end of the previous period ( $t-1$ ). Following

DeMiguel et al. (2009a) and Kirby and Ostdiek (2012), the proportionate transaction cost ( $T_i$ ) of trading US individual equities and equity indices is set at 50 bps. We assume that transaction costs are a linear function of trade value, the same for buys and sells; and we only take account of the variable costs of trading assets.

We compare the performance of MV and  $1/N$  applied to each of the stages of a two-stage process, and this can be done in four different ways:-

- 1 Apply MV to the ten asset classes, and then  $1/N$  to the individual assets in each asset class.
- 2 Apply  $1/N$  to the ten asset classes, and then  $1/N$  to the individual assets in each asset class.
- 3 Apply MV to the ten asset classes, and then MV to the individual assets in each asset class.
- 4 Apply  $1/N$  to the ten asset classes, and then MV to the individual assets in each asset class.

For the reasons given above, we expect that using MV for asset allocation, and then  $1/N$  for stock selection is superior to the other three two-stage procedures, i.e. procedure 1 is dominant. For similar reasons, we also expect that the opposite strategy of using  $1/N$  for asset allocation, and then MV for stock selection is inferior to the other three procedures, i.e. procedure 4 is dominated. Having computed the portfolio asset weights for the first and second stages using MV or  $1/N$ , we use these weights to calculate the resulting weights of the 93 companies in the ten industrial sectors for each of the four two-stage procedures. The out-of-sample performance of the four portfolios is then compared using seven different performance measures.

## 5. Results

In stage one we use both  $1/N$  and MV to form a portfolio of the ten industrial sectors, i.e. asset allocation. In Table 1 we compare the out-of-sample performance using the Sharpe, Dowd, and Omega ratios, and certain equivalent returns<sup>8</sup> (CER). (See Appendix 2 for details of the performance measures.)

This shows that MV is superior to  $1/N$  on every performance measure for asset allocation.

Performance Measure	$1/N$	MV	Significant Difference ( $\rho$ )
CER	0.0613	0.0712	0.1861
Sharpe Ratio	0.6190	0.6939	0.1824
Dowd Ratio	0.4019	0.4779	-
Omega Ratio	1.7999	1.9285	-

Table 1: Performance of  $1/N$  and MV in Forming Portfolios of the Ten Industrial Sectors, i.e. Stage One Asset Allocation - 12-month expanding estimation window, January 1994 to August 2017,  $\lambda = 5$ , Transaction costs = 50 bps.

---

<sup>8</sup> We assume that investors have a time-separable CRRA utility function and unit initial wealth.

We tested for significant differences between the Sharpe ratios of these four two-stage techniques. Following DeMiguel et al. (2009a), we used the approach suggested by Jobson and Korkie (1980), as corrected by Memmel (2003). Given two portfolios  $i$  and  $n$ , with  $\hat{\mu}_i, \hat{\mu}_n, \hat{\sigma}_i, \hat{\sigma}_n$  and  $\hat{\sigma}_{i,n}$  as their estimated means, variances, and covariance over a sample of size  $T-M$ , the test of the hypothesis  $H_0 : \hat{\mu}_i / \hat{\sigma}_i = \hat{\mu}_n / \hat{\sigma}_n$  is obtained by the test statistic  $Z$ , which is asymptotically distributed as a standard

normal:

$$Z_{JK} = (\hat{\sigma}_n \hat{\mu}_i - \hat{\sigma}_i \hat{\mu}_n) / \sqrt{\mathcal{G}}$$

where  $\mathcal{G} = (2\hat{\sigma}_i^2 \hat{\sigma}_n^2 - 2\hat{\sigma}_i \hat{\sigma}_n \hat{\sigma}_{i,n} + \hat{\mu}_i^2 \hat{\sigma}_n^2 / 2 - \hat{\mu}_i \hat{\mu}_n \hat{\sigma}_{i,n}^2 / \hat{\sigma}_i \hat{\sigma}_n) / (T-M)$

Consistent with DeMiguel et al. (2009a), we used the following method to test the significance of differences between the CERs. If  $\hat{\nu}$  denotes the vector of moments,  $\nu = (\mu_i, \mu_n, \hat{\sigma}_i^2, \hat{\sigma}_n^2)$ , then  $\hat{\nu}$ , its empirical counterpart, is obtained from a sample of size  $T-M$ . The difference between the CERs of the two strategies  $i$  and  $n$ ,  $f(\nu) = (\mu_i - \gamma \hat{\sigma}_i^2 / 2) - (\mu_n - \gamma \hat{\sigma}_n^2 / 2)$  is the asymptotic distribution of

$$f(\nu) \text{ is } \sqrt{T}(f(\hat{\nu}) - f(\nu)) \rightarrow N\left(0, \frac{\partial f^T}{\partial \nu} \Theta \frac{\partial f}{\partial \nu}\right), \text{ where } \Theta = \begin{pmatrix} \sigma_i^2 & \sigma_{i,n} & 0 & 0 \\ \sigma_{i,n} & \sigma_n^2 & 0 & 0 \\ 0 & 0 & 2\sigma_i^4 & 2\sigma_{i,n}^2 \\ 0 & 0 & 2\sigma_{i,n}^2 & 2\sigma_n^4 \end{pmatrix}$$

The resulting probability values appear in Table 1, and show that the differences between  $1/N$  and MV are only significant at the 18% level.

In stage two we use both  $1/N$  and MV to form ten portfolios of individual companies for each of the industrial sectors. The out-of-sample performance of these 20 portfolios appears in Table 2 measured in four different ways. For six of the ten industrial sectors  $1/N$  is superior to MV on all four performance measures, and for another two sectors  $1/N$  is superior to MV on three of the measures. These results indicate that, across the ten sectors,  $1/N$  is generally superior to MV at stock selection. The two final



columns of Table 2 show the probabilities of significance tests on the differences between CERs and Sharpe ratios for the  $1/N$  and MV portfolios. While only four of the CERs, and two of the Sharpe ratios are significantly different, the average significance level across all the CERs is 0.153, and for the Sharpe ratios it is 0.210. This suggests that  $1/N$  tends to outperform MV, but not at a conventional level of significance.

Industrial Sectors	CER		Sharpe Ratio		Dowd Ratio		Omega Ratio		Sig. Differences	
	1/N	MV	1/N	MV	1/N	MV	1/N	MV	CER <sup>†</sup>	SR
Consumer Discretionary	0.0287	0.0451	0.5505	0.5604	0.3300	0.3427	1.6863	1.7099	0.23038	0.46442
Consumer Staples	0.0802	0.0727	0.7680	0.7050	0.5591	0.4890	2.0650	1.9403	0.16307	0.1519
Energy	-0.0063	0.0274	0.4614	0.4815	0.2608	0.2797	1.5321	1.5786	0.0554*	0.40131
Financials	-0.0290	-0.0062	0.4652	0.4129	0.2619	0.2281	1.5969	1.5417	0.17236	0.2678
Health Care	0.0908	0.0627	0.8223	0.6382	0.6062	0.4152	2.0631	1.8131	0.0062***	0.0061***
Industrials	0.0561	0.0388	0.6395	0.5236	0.4089	0.3131	1.7792	1.6463	0.0978*	0.0546*
Information Technology	0.0057	0.0172	0.5303	0.4894	0.3115	0.2832	1.6249	1.5834	0.3007	0.31945
Materials	0.0347	0.0265	0.5479	0.4589	0.3295	0.2639	1.6819	1.5751	0.31348	0.15538
Telecommunications	-0.0034	-0.0083	0.3636	0.3493	0.1966	0.1873	1.4594	1.4427	0.0535*	0.15717
Utilities	0.0639	0.0544	0.6461	0.5789	0.4224	0.3635	1.8064	1.7285	0.13984	0.1234

Table 2: Performance of  $1/N$  and MV in Forming Portfolios of the Shares Within Each Industrial Sector, i.e. Stage Two Stock Selection - 12-month expanding estimation window, January 1994 to August 2017,  $\lambda = 5$ , Transaction costs = 50 bps.

Significance at the 10%, 5% and 1% levels is denoted by \*, \*\* and \*\*\* respectively

† The significance tests for the CERs use the value of  $\lambda$  used in forming the portfolios.

We now use the stage one results, which allocate money between sectors using either  $1/N$  or MV, and the stage two results, which select individual companies within each of these sectors using either  $1/N$  or MV, to compute the overall portfolio of 93 companies. The out-of-sample performance of the four ways of forming the overall portfolio is presented in Table 3.

Performance Measure	MV-1/N	1/N-1/N	MV-MV	1/N-MV
CER	0.0925	0.0818	0.0745	0.0719
Sharpe Ratio	0.8614	0.7673	0.7248	0.7002
Dowd Ratio	0.6706	0.5498	0.5132	0.4854
Omega Ratio	2.2252	2.0213	2.0206	1.9331

Table 3: Performance Measures for the Four Two-stage Procedures for the US Data - 12-month Expanding Estimation Window, January 1994 to August 2017,  $\lambda = 5$ , Transaction costs = 50 bps.

On each of our four performance measures MV-1/N is the best, and the reverse procedure of 1/N-MV

is the worst. These results support our hypothesis that the two-stage procedure of MV, followed by  $1/N$  is superior to the other three procedures. The pure strategies rank in the middle, with  $1/N$ - $1/N$  superior on six of the measures to MV-MV. Figure 1 plots the Sharpe ratios<sup>9</sup> of the best and worst two-stage procedures as we add additional out-of-sample periods. These show that, in accordance with our hypothesis, MV- $1/N$  is consistently superior to  $1/N$ -MV.

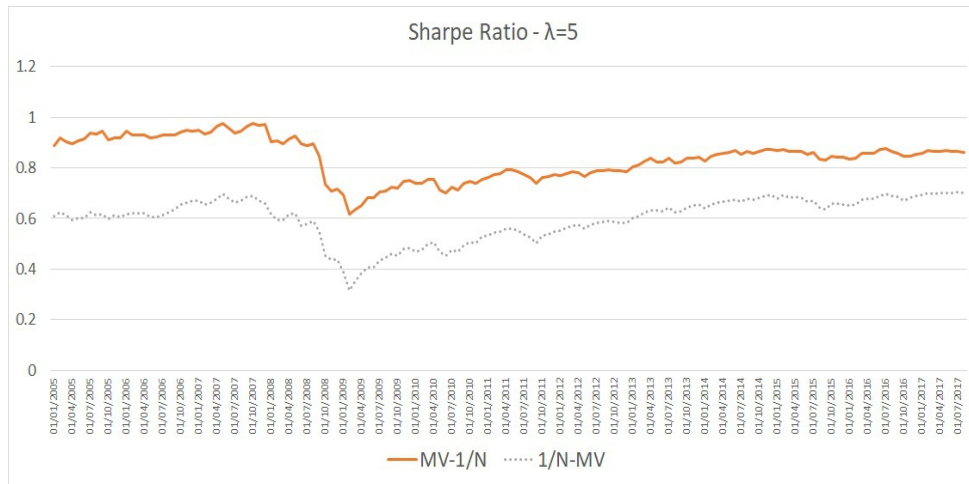


Figure 1: Time Series Plots of the Sharpe Ratios Comparing MV- $1/N$  with  $1/N$ -MV.  $\lambda = 5$

The probability values of significance tests of differences between the four two-stage procedures appear in Table 4, with the implied quasi-ordering in Figure 2 using the three significant relationships. These show that MV- $1/N$  is significantly superior to both MV-MV and  $1/N$ -MV, and that  $1/N$ - $1/N$  is also significantly superior to  $1/N$ -MV using both Sharpe ratios and CERs. The significance tests in Tables 1, 2 and 3 show that the benefits of using MV in the first stage and  $1/N$  in the second stage are modest, but over the two stages the benefits cumulate; and so the combined MV- $1/N$  two-stage investment process has statistically significant benefits. This is consistent with the common finding that tests of  $1/N$  versus MV in a one stage process are often not as clear cut as our two stage results.

<sup>9</sup> The corresponding plots for the CERs are virtually identical to those for the Sharpe ratios.

Superior	Inferior	Sharpe Ratio	CER <sup>†</sup>
MV-1/N	1/N-MV	0.029**	0.026**
MV-1/N	MV-MV	0.007***	0.005***
1/N-1/N	1/N-MV	0.087*	0.085*
MV-1/N	1/N-1/N	0.155	0.198
1/N-1/N	MV-MV	0.361	0.323
MV-MV	1/N-MV	0.388	0.401

Table 4: Probability Values for the Significance Tests of the Differences Between Sharpe Ratios and CERs for the Four Two-Stage Procedures.  $\lambda = 5$ .

Significance at the 10%, 5% and 1% levels is denoted by \*, \*\* and \*\*\* respectively

† The significance tests for the CERs use the value of  $\lambda$  used in forming the portfolios.

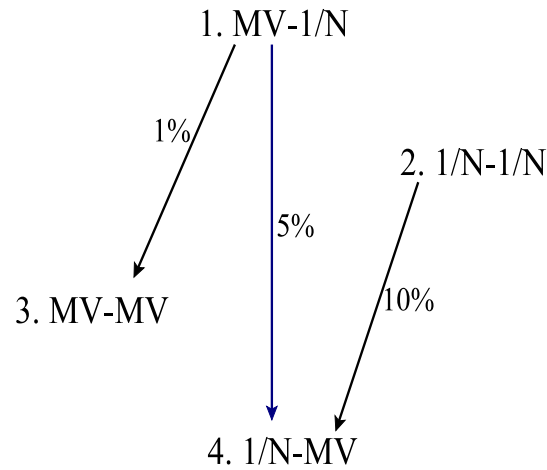


Figure 2: Quasi-Ordering of the Four Two-stage Procedures Based on the Statistically Significant Differences in Sharpe Ratios and CERs in Table 4,  $\lambda = 5$

## 6. Simulation Results

Based on parameters calibrated to individual US stocks, DeMiguel et al. (2009a) argue that, given the limited history of stock markets, the length of the estimation window needed for MV and its extensions to outperform  $1/N$  is too large to obtain empirical results. They do not consider forming portfolios using asset classes because they believe the lower idiosyncratic volatility of asset classes leads to a higher utility loss than when  $1/N$  is used. However, we argue that this is not the case, and the lower idiosyncratic volatility of asset classes improves the performance of MV, increasing the likelihood of MV

outperforming  $1/N$ .

#### 6.A. DeMiguel et al. (2009a)'s Relative Performance Measure

In this section we explain DeMiguel et al. (2009a)'s performance measure which we use to evaluate the relative performance of portfolio selection techniques. Following DeMiguel et al. (2009a), we assume a one factor structure for excess returns:-

$$\mathbf{R}_{a,t} = \boldsymbol{\alpha} + \boldsymbol{\beta}R_{b,t} + \boldsymbol{\varepsilon}_t$$

where  $\mathbf{R}_{a,t}$  is the excess asset returns vector,  $\boldsymbol{\alpha}$  is the vector of mispricing coefficients,  $\boldsymbol{\beta}$  is the factor loadings vector,  $R_{b,t}$  is the excess return on the factor portfolio where  $R_b \sim N(\mu_b, \sigma_b)$ , and  $\boldsymbol{\varepsilon}_t$  is the noise vector,  $\boldsymbol{\varepsilon} \sim N(0, \boldsymbol{\Sigma}_\varepsilon)$ , which is independent from the factor portfolio. Given this structure, the expected excess returns vector is  $\boldsymbol{\mu}_\alpha = \boldsymbol{\alpha} + \boldsymbol{\beta}\mu_b$ , and its variance-covariance matrix is  $\boldsymbol{\Sigma}_\alpha = \boldsymbol{\beta}\sigma_b^2\boldsymbol{\beta}^T + \boldsymbol{\Sigma}_\varepsilon$ .

The expected loss function defined in Kan and Zhou (2007) and DeMiguel et al. (2009a) is given by:

$$L(\mathbf{x}^*, \mathbf{x}) = U(\mathbf{x}^*) - E[U(\mathbf{x})]$$

where  $\mathbf{x}^*$  is a vector of the optimal portfolio weights with  $\mathbf{x}^* = \boldsymbol{\Sigma}_\alpha^{-1}\boldsymbol{\mu}_\alpha/\lambda$ ,  $\lambda$  is the investor's risk aversion parameter, and  $U(\mathbf{x})$  is the utility of a portfolio with weights given by  $\mathbf{x}$ , where:-

$$U(\mathbf{x}) = \mathbf{x}^T \boldsymbol{\mu}_\alpha - \frac{\lambda}{2} \mathbf{x}^T \boldsymbol{\Sigma}_\alpha \mathbf{x} = \mathbf{x}^T (\boldsymbol{\alpha} + \boldsymbol{\beta}\mu_b) - \frac{\lambda}{2} \mathbf{x}^T (\boldsymbol{\beta}\sigma_b^2\boldsymbol{\beta}^T + \boldsymbol{\Sigma}_\varepsilon) \mathbf{x}, \quad (9)$$

and

$$U(\mathbf{x}^*) = \frac{\boldsymbol{\mu}_\alpha^T \boldsymbol{\Sigma}_\alpha^{-1} \boldsymbol{\mu}_\alpha}{2\lambda} = \frac{(\boldsymbol{\alpha} + \boldsymbol{\beta}\mu_b)^T (\boldsymbol{\beta}\sigma_b^2\boldsymbol{\beta}^T + \boldsymbol{\Sigma}_\varepsilon)^{-1} (\boldsymbol{\alpha} + \boldsymbol{\beta}\mu_b)}{2\lambda}$$

Since our goal is to compare the performance of sample-based MV and its extensions with that of  $1/N$ , we focus on the loss difference between  $L(\mathbf{x}^*, \mathbf{x}^{ew})$  and  $L(\mathbf{x}^*, \hat{\mathbf{x}})$ , which is defined as:-

$$\Gamma(\hat{\mathbf{x}}, \mathbf{x}^{ew}) = 2\lambda\{E[U(\hat{\mathbf{x}})] - U(\mathbf{x}^{ew})\}$$

where  $\hat{\mathbf{x}}$  is the vector of portfolio weights given by an optimisation model,  $\mathbf{x}^{ew}$  is the equally weighted policy ( $1/N$ ), and:-

$$U(\mathbf{x}^{ew}) = \frac{(\mathbf{1}_N^\top \boldsymbol{\mu}_a)^2}{2\lambda \mathbf{1}_N^\top \boldsymbol{\Sigma}_a \mathbf{1}_N} = \frac{\Phi(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_\epsilon; \mu_b, \sigma_b^2)}{2\lambda}$$

where

$$\Phi(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_\epsilon; \mu_b, \sigma_b^2) = \frac{[\mathbf{1}_N^\top (\boldsymbol{\alpha} + \boldsymbol{\beta} \mu_b)]^2}{\mathbf{1}_N^\top (\boldsymbol{\beta} \sigma_b^2 \boldsymbol{\beta}^\top + \boldsymbol{\Sigma}_\epsilon) \mathbf{1}_N}.$$

see the derivation in Appendix B in DeMiguel et al. (2009a). We focus on  $\hat{\mathbf{x}}$  computed using MV with sample means  $\hat{\boldsymbol{\mu}}_a$  and covariances  $\hat{\boldsymbol{\Sigma}}_a$ , i.e.  $\hat{\mathbf{x}} = \frac{\hat{\boldsymbol{\Sigma}}_a^{-1} \hat{\boldsymbol{\mu}}_a}{\lambda}$  which facilitates the comparison of our simulation results with those of DeMiguel et al. (2009a).

As shown by Kan and Zhou (2007), when  $\hat{\mathbf{x}} = \frac{\hat{\boldsymbol{\Sigma}}_a^{-1} \hat{\boldsymbol{\mu}}_a}{\lambda}$ , and assuming the distribution of returns is jointly normal, the expected utility is given by:-

$$\mathbb{E}[U(\hat{\mathbf{x}})] = \frac{k}{2\lambda} \Theta(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_\epsilon; \mu_b, \sigma_b^2) - \frac{h}{2\lambda}$$

where

$$k = \frac{T}{T - N - 2} \left[ 2 - \frac{T^2 - 2T}{(T - N - 1)(T - N - 4)} \right]$$

and

$$h = \frac{TN(T - 2)}{(T - N - 1)(T - N - 2)(T - N - 4)}$$

and

$$\Theta(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_\epsilon; \mu_b, \sigma_b^2) = (\boldsymbol{\alpha} + \boldsymbol{\beta} \mu_b)^\top (\boldsymbol{\beta} \sigma_b^2 \boldsymbol{\beta}^\top + \boldsymbol{\Sigma}_\epsilon)^{-1} (\boldsymbol{\alpha} + \boldsymbol{\beta} \mu_b)$$

and  $T$  is the number of months (the size of the sample), and  $N$  is the number of assets. Therefore:-

$$\Gamma(\hat{\mathbf{x}}, \mathbf{x}^{ew}) = k\Theta(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_\epsilon; \mu_b, \sigma_b^2) - \Phi(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_\epsilon; \mu_b, \sigma_b^2) - h.$$

If  $\Gamma(\hat{\mathbf{x}}, \mathbf{x}^{ew}) > 0$ , the MV strategy outperforms  $1/N$ ; otherwise  $1/N$  outperforms MV. In the next subsection we use this performance measure for a simulation analysis of MV and  $1/N$ .

### 6.B. Simulation with Non-Zero $\alpha$ s

Since we have monthly data, the variables and parameters in our simulations are set to a monthly

frequency. To be comparable with the simulations of DeMiguel et al. (2009a), all our settings are consistent with theirs, except for the  $\alpha$ s. Our  $\beta$ s are drawn uniformly from the 0.5 to 1.5 range using a market factor with monthly excess returns with a mean of  $\mu_b = 0.67\%$ , and a variance of  $\sigma_b^2 = 0.21\%$ , as assumed by DeMiguel et al. 2009a. We also follow DeMiguel et al. (2009a) in assuming the covariance matrix is diagonal. DeMiguel et al. (2009a) assume the  $\alpha$ s are zero, but this assumption is too strong, as zero  $\alpha$ s only exist in perfect equilibrium. Although the estimated  $\alpha$ s are often statistically insignificant from zero, they are never exactly zero. As pointed out by Jarrow (2010), the conditions for a strictly zero alpha are so strong that “in the complexity of actual markets, their violation, although anathema to economists, is not logically unreasonable.” Therefore we follow a more realistic approach, and draw the  $\alpha$ s from a normal distribution with zero mean and a standard deviation of 30 bps, which matches the cross-sectional standard deviation of the  $\alpha$ s of the ten stocks in the Financials sector. The annualised idiosyncratic volatilities vary from 2% to 26%. To compare our results with those of DeMiguel et al. (2009a) we conduct two simulations - first, the  $\alpha$ s are strictly zero, as assumed by DeMiguel et al. (2009a); and second, the  $\alpha$ s have zero mean but a non-zero standard deviation. We compute the 95% confidence intervals based on the simulation results.

The plots of the simulated values of  $\Gamma(\hat{\mathbf{x}}, \mathbf{x}^{ew})$  against average idiosyncratic volatility ( $\sigma_\epsilon$ ) are shown in Figure 3. These results are based on 1,000 simulations in which  $T = 240$  and  $N = 10$ . The upper panel of Figure 3 shows the scenario depicted in DeMiguel et al. (2009a) with  $\alpha = 0$ . They write that “all else being equal, the performance of the sample-based mean-variance (and that of the optimizing policies in general) would improve relative to that of the  $1/N$  policy if the idiosyncratic asset volatility was much higher than 20%.” However, the lower panel of Figure 3 shows that DeMiguel et al. (2009a)’s conclusion about idiosyncratic volatility is not robust to allowing the  $\alpha$ s to be stochastic. When the  $\alpha$ s are allowed to vary across assets, even a small cross-sectional variation in the  $\alpha$ s (30 bps standard deviation centred around zero), reverses the results of DeMiguel et al. (2009a). The smaller is idiosyncratic asset volatility ( $\sigma_\epsilon$ ), the better MV performs relative to  $1/N$ . All the values of  $\Gamma(\hat{\mathbf{x}}, \mathbf{x}^{ew})$  in the upper panel of Figure 3 are

negative, and so  $1/N$  is preferred to MV for all levels of idiosyncratic volatility. In the lower panel of Figure 3 the values of  $\Gamma(\hat{\mathbf{x}}, \mathbf{x}^{ew})$  are similar at approximately zero, but for volatility below about 10%, MV becomes progressively superior to  $1/N$ . Thus given our simulation setting, as long as the average annualised idiosyncratic volatility ( $\sigma_\epsilon$ ) is lower than about 10%, even with just 240 observations, MV outperforms  $1/N$ . Of course, this also depends on other factors.

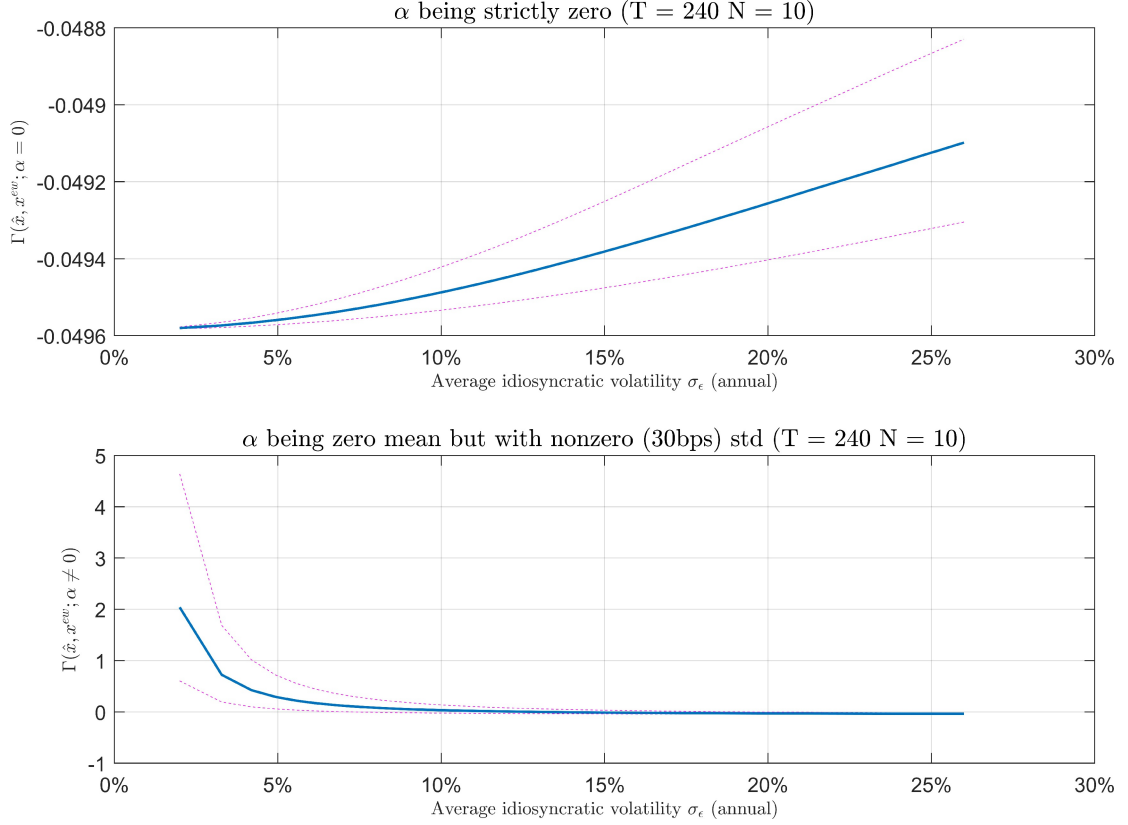


Figure 3:  $\Gamma(\hat{\mathbf{x}}, \mathbf{x}^{ew})$  Versus Idiosyncratic Volatility ( $\sigma_\epsilon$ ). The simulation is based on  $T = 240$  and  $N = 10$ . The dashed lines are 95% confidence intervals.

To have a clearer idea about how changes in the  $\alpha$ s can affect  $\Gamma(\hat{\mathbf{x}}, \mathbf{x}^{ew})$ , in Figure 4 we plot  $\Gamma(\hat{\mathbf{x}}, \mathbf{x}^{ew})$  against the cross-sectional standard deviation of the  $\alpha$ s ( $\sigma_\alpha$ ), as well as the average idiosyncratic volatility ( $\sigma_\epsilon$ ). The larger the sample size, the more likely is MV to outperform  $1/N$ ; and so we use  $T = 360$  in this simulation to produce insights useful for benchmarking the simulation results below in Section 6.C using real world parameters. We can see that when the  $\alpha$ s are more dispersed and the idiosyncratic volatilities ( $\sigma_\epsilon$ ) are smaller, MV and its extensions are more likely to outperform  $1/N$ . In

the next section, we calibrate our simulation to our monthly return data for both individual stocks and the ten sector indices to give a more realistic and comprehensive performance comparison.

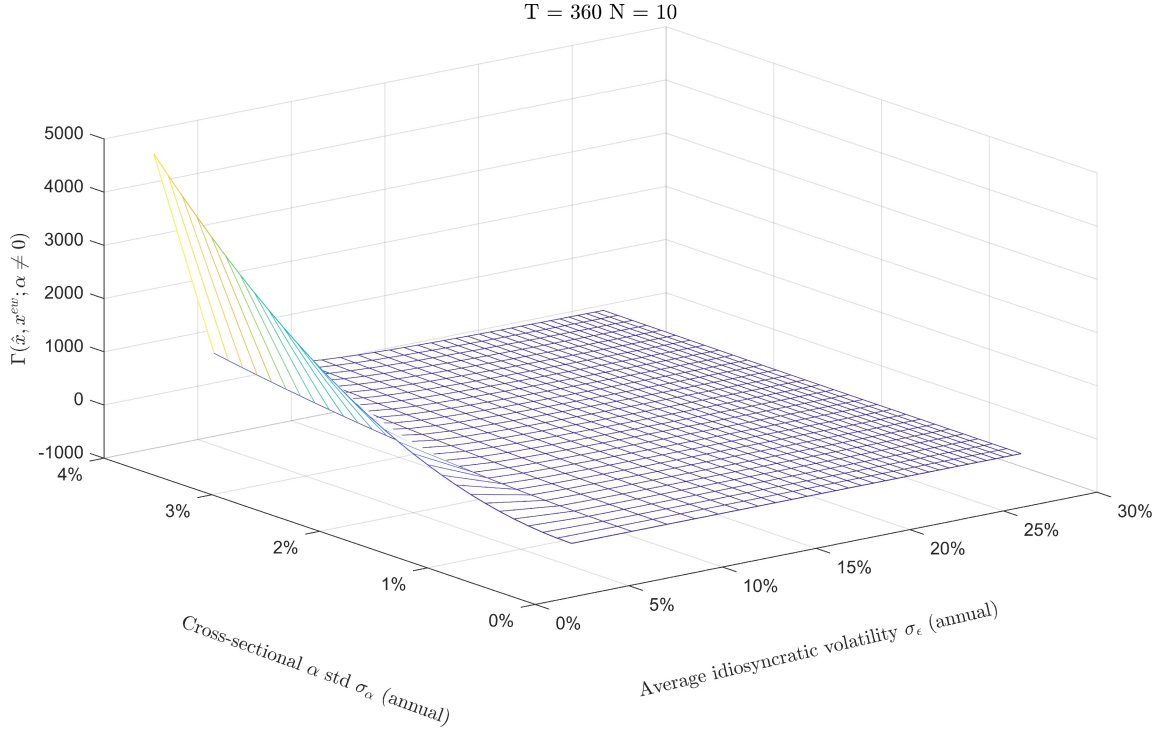


Figure 4.  $\Gamma(\hat{\mathbf{x}}, \mathbf{x}^{ew})$  Versus Average Cross-sectional Idiosyncratic Volatility ( $\sigma_\epsilon$ ) and the Cross-Sectional Standard Deviation of  $\alpha$  ( $\sigma_\alpha$ ). The simulation is based on  $T = 360$  and  $N = 10$ .

#### 6.C. Simulation Based on Real World Parameters

To gain more insight into when MV and its extensions outperform  $1/N$ , we calibrate the simulation to match our monthly return data on ten sector indices and 93 companies, and continue to use  $\Gamma(\hat{\mathbf{x}}, \mathbf{x}^{ew})$  as a proxy for the relative performance of MV versus  $1/N$ . A positive value of  $\Gamma(\hat{\mathbf{x}}, \mathbf{x}^{ew})$  can be regarded as the likelihood of MV outperforming  $1/N$ , and the higher this value, the better is the relative performance of MV. If the value of  $\Gamma(\hat{\mathbf{x}}, \mathbf{x}^{ew})$  is higher than 50%, we say MV outperforms  $1/N$ .

To gain a broader sense of the relative performance of extensions to MV, we also consider the Bayes-Stein model, in addition to the Markowitz sample-based MV model. For Bayes-Stein the vector of

portfolio weights is  $\hat{\mathbf{x}} = \frac{\widehat{\Sigma}^{BS}{}^{-1} \widehat{\boldsymbol{\mu}}^{BS}}{\lambda} \mathbf{a}$  where  $\widehat{\boldsymbol{\mu}}^{BS}$  and  $\widehat{\Sigma}^{BS}$  are the Bayes-Stein shrinkage estimates of the



means and covariance matrix. When Bayes-Stein is used  $\mathbb{E}[U(\hat{\mathbf{x}})]$  no longer has a closed form expression. Therefore, we compute the simulated values of  $U(\hat{\mathbf{x}})$  (250 values are simulated based on eqn. 1) and use the resulting average to approximate  $\mathbb{E}[U(\hat{\mathbf{x}})]$ .

In this simulation we use the S&P500 index as our single factor whose sample monthly excess return ( $\mu_b$ ) has a mean of 0.64%, and a variance ( $\sigma_b^2$ ) of 0.18%. We consider 11 portfolios. The first ten consist of the constituent stocks in the ten industrial sectors (see Appendix 1), i.e the second stage stock selection problem. The last portfolio includes the ten sector indices, i.e. the first stage asset allocation problem. The values of  $\alpha$  and  $\beta$  are drawn from a uniform distribution matching the first two moments of the sample cross-sectional distribution of  $\alpha$  and  $\beta$  in the corresponding portfolio<sup>10</sup>. Consistent with DeMiguel et al. (2009a),  $\Sigma_\epsilon$  is a diagonal matrix, with its diagonal elements drawn from a uniform distribution matching the minimum and maximum of the cross-sectional idiosyncratic variance ( $\sigma_\epsilon$ ) of the corresponding portfolio. Using the full sample, we estimate the idiosyncratic asset volatilities of the shares available for inclusion in the ten sector portfolios and the ten sector indices. Table 5 summarises the values of  $\Gamma(\hat{\mathbf{x}}, \mathbf{x}^{ew})$  for both the Markowitz and Bayes-Stein models computed from 1,000 simulations.

From Table 5 we see that in these 11 portfolios, all the average annualised idiosyncratic volatilities ( $\sigma_\epsilon$ ) of portfolios of individual equities, except for SP5ECST (consumer staples) at 18.9%, are higher than or equal to 20%. Only the portfolio of sector indices has a much lower average annualised idiosyncratic volatility ( $\sigma_\epsilon$ ) of 11.9%. The positive values of  $\Gamma(\hat{\mathbf{x}}, \mathbf{x}^{ew})$  show that Markowitz is clearly inferior to  $1/N$  for stage two stock selection as most  $\Gamma(\hat{\mathbf{x}}, \mathbf{x}^{ew})$  values are zero, and the highest is only 22.5%. However for first stage asset allocation, when  $T = 360$   $\Gamma(\hat{\mathbf{x}}, \mathbf{x}^{ew})$  is 58.8%, and Markowitz is superior to  $1/N$ . These results are consistent with Figure 5. The 11.9% average annual idiosyncratic volatility ( $\sigma_\epsilon$ ) of the indices is much smaller than that of the stock portfolios, but still higher than the 10% critical level mentioned

---

<sup>10</sup> We also assumed normal distributions for the values of  $\alpha$  and  $\beta$ , and the results are similar to those for uniform distributions.

in Section 6.B. When  $T = 240$ , the positive rate for indices using Markowitz is only 11.6%; and only when  $T = 360$  do we observe the much higher positive rate of 58.8%. This is because a larger sample leads to better estimates of the MV input parameters.

Portfolio	N	Markowitz % Positive Values of $\Gamma(\hat{\mathbf{x}}, \mathbf{x}^{ew})$			Bayes-Stein % Positive Values of $\Gamma(\hat{\mathbf{x}}, \mathbf{x}^{ew})$			Idiosyncratic Volatility ( $\sigma_e$ ) (annual)			Alpha (annual)			Beta		
		T = 120	T = 240	T = 360	T = 120	T = 240	T = 360	min	max	mean	min	max	mean	min	max	mean
SP5EFIN	10	0.000	0.000	0.000	0.000	0.000	0.038	20.3%	64.4%	28.8%	-6.1%	5.8%	1.1%	0.95	1.92	1.39
SP5ETEL	3	0.000	0.000	0.000	0.000	0.000	0.000	20.5%	24.3%	21.7%	0.7%	1.9%	1.4%	0.66	0.84	0.73
SP5EHCR	10	0.000	0.000	0.001	0.000	0.023	0.233	16.2%	29.4%	22.8%	4.5%	11.9%	7.6%	0.51	0.91	0.68
SP5ECST	10	0.000	0.000	0.025	0.000	0.097	0.453	16.2%	23.7%	18.9%	2.8%	13.6%	5.7%	0.23	0.65	0.47
SP5EIND	10	0.000	0.000	0.000	0.000	0.007	0.158	16.3%	24.9%	21.4%	-1.8%	6.2%	3.8%	0.62	1.30	0.98
SP5EMAT	10	0.000	0.000	0.000	0.000	0.007	0.106	17.4%	40.9%	24.4%	-3.0%	8.4%	3.0%	0.38	1.45	1.01
SP5ECOD	10	0.000	0.002	0.066	0.000	0.143	0.492	16.8%	39.0%	26.4%	-6.8%	6.7%	2.4%	0.63	1.79	1.14
SP5EUTL	10	0.000	0.000	0.000	0.000	0.000	0.017	16.6%	24.8%	20.0%	4.9%	9.3%	6.7%	0.08	0.39	0.29
SP5EENE	10	0.000	0.000	0.004	0.000	0.028	0.230	14.6%	33.5%	25.9%	0.8%	10.0%	3.5%	0.55	1.32	0.93
SP5EINT	10	0.000	0.008	0.225	0.003	0.339	0.761	16.9%	39.4%	28.9%	-5.6%	10.9%	3.3%	0.81	1.58	1.33
Indices	10	0.000	0.116	0.588	0.029	0.628	0.892	7.4%	15.2%	11.9%	-1.7%	5.2%	1.2%	0.42	1.43	0.91

Table 5: Positive Values of  $\Gamma(\hat{\mathbf{x}}, \mathbf{x}^{ew})$  for Markowitz and Bayes-Stein, and Simulation Characteristics

The results for Bayes-Stein in Table 5 reinforce the view that ‘shrinkage estimators’ are superior to sample-based MV estimates in the portfolio optimisation context (Jobson & Korkie, 1980), and provide stronger support for our hypothesis that MV is superior for stage one, and  $1/N$  is superior for stage two. Even with only 240 observations, Bayes-Stein has a  $\Gamma(\hat{\mathbf{x}}, \mathbf{x}^{ew})$  value of 62.8% for the index portfolio; and with 360 observations  $\Gamma(\hat{\mathbf{x}}, \mathbf{x}^{ew})$  is close to 90%, indicating that MV is clearly superior to  $1/N$  for first stage asset allocation. For the second stage of stock selection, the only sector for which Bayes-Stein is superior to  $1/N$  is information technology (SP5INT) with a  $\Gamma(\hat{\mathbf{x}}, \mathbf{x}^{ew})$  value of 76.1%. Although the average idiosyncratic volatility in this sector is high (around 29%), the  $\alpha$ s in this sector are distributed across a relatively wide range (16.5%). This observation is consistent with the results in Figure 4, in that the cross-sectional standard deviation of the  $\alpha$ s has a positive influence on Bayes-Stein’s performance, relative to  $1/N$ .

These simulation results, calibrated to the real world, are consistent with our empirical results in Section

5, and support our hypothesis that MV is superior for asset allocation, and  $1/N$  is superior for stock selection. They also show that relaxing DeMiguel et al. (2009a)'s assumption of zero  $\alpha$ s reverses their conclusions concerning the effect of idiosyncratic volatility on the relative performance of MV and  $1/N$ .

## 7. Robustness Checks

We now run a set of robustness checks on our core results presented in Section 5 and vary the choice of the optimal diversification model and the various parameters involved. We use the -Litterman (1992) and the Bayesian diffuse prior models as alternatives to Bayes-Stein. Black-Litterman is a popular portfolio optimization method for dealing with estimation risk, and has recently attracted considerable interest in both academia and the financial services industry, e.g. Bessler and Wolff (2017), Kolm et al (2014), Oikonomou et al. (2018), Platanakis and Sutcliffe (2017), Platanakis et al. (2019), Platanakis and Urquhart (2019) and Silva, et al. (2017). We also use the Markowitz and Bayesian diffuse prior (BDP) models.

*Black-Litterman.* This model combines two sources of information: the investors' 'views' on asset returns (subjective return estimates), and the reference (or benchmark) portfolio used for estimating the 'neutral' (or 'implied') return estimates. The original Black-Litterman model computes the column vector of implied excess-returns, denoted by  $\mathbf{H}$ , as follows:-

$$\mathbf{H} = \lambda \Sigma \mathbf{x}^{\text{Reference}}, \quad (10)$$

where  $\mathbf{x}^{\text{Reference}}$  is a column vector with the weights of the benchmark (reference) portfolio. We set  $\mathbf{x}^{\text{Reference}}$  to the global minimum-variance portfolio, as in Bessler et al. (2017) and Platanakis et al. (2019), amongst others, which represents the case where investors think asset return estimates are highly exposed to estimation risk.

The Black-Litterman model computes the column vector of mean returns as follows:-

$$\boldsymbol{\mu}_{BL} = \left[ (c\boldsymbol{\Sigma})^{-1} + \mathbf{P}^T \boldsymbol{\Omega}^{-1} \mathbf{P} \right]^{-1} \left[ (c\boldsymbol{\Sigma})^{-1} \mathbf{H} + \mathbf{P}^T \boldsymbol{\Omega}^{-1} \mathbf{Q} \right], \quad (11)$$

where  $\mathbf{P}$  is a binary diagonal matrix with a value of unity if a subjective return estimate exists for the asset. The column vector  $\mathbf{Q}$  contains the investor's views (subjective return estimates), and the parameter  $c$  is a measure of the reliability of the implied excess-returns in  $\mathbf{H}$ , which we set to 0.1625 as in Bessler and Wolff (2017), Platanakis and Sutcliffe (2017) and Platanakis et al (2019), amongst others. We follow Meucci (2010) and many others, and compute the diagonal matrix  $\boldsymbol{\Omega}$  as follows:-

$$\boldsymbol{\Omega} = \frac{1}{\delta} \mathbf{P} \boldsymbol{\Sigma} \mathbf{P}^T. \quad (12)$$

Following Meucci (2010) and Platanakis and Sutcliffe (2017), amongst others, we set  $1/\delta$  to unity, and use mean asset returns as the subjective return estimates in the column vector  $\mathbf{Q}$ . We follow Satchell and Scowcroft (2000) and many other studies, and compute the posterior covariance matrix ( $\boldsymbol{\Sigma}_{BL}$ ) as follows:-

$$\boldsymbol{\Sigma}_{BL} = \boldsymbol{\Sigma} + \left[ (c\boldsymbol{\Sigma})^{-1} + \mathbf{P}^T \boldsymbol{\Omega}^{-1} \mathbf{P} \right]^{-1} \quad (13)$$

*Markowitz.* The Markowitz (1952) model with VBCs appears in equation 7, where the inputs are historic values.

*Bayesian Diffuse Prior.* Barry (1974) and Klein and Bawa (1976), among many others, have shown that the predictive distribution of asset returns has the following estimates for the mean returns and covariance matrix (student's  $t$ -distribution):-  $(\boldsymbol{\mu}, (1+1/T)\boldsymbol{\Sigma})$  subject to a diffuse prior  $(p(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-(N+1)/2})$  and a normal conditional likelihood. As a result, this portfolio construction

technique inflates the sample covariance matrix by a factor of  $(1+1/T)$ .

We varied five aspects of the analysis - the MV portfolio model (Bayes-Stein, Black-Litterman, Markowitz or Bayesian diffuse prior), the risk aversion parameter, ( $\lambda = 2, 5$  or  $10$ ), the VBC parameter ( $\alpha = 0.10, 0.15$  and  $0.20$ ), the initial length of the expanding window (12, 36 and 60 months), and the type of constraint (VBC or UGC<sup>11</sup>). Table 6 shows the CERs and Sharpe ratios for the four robustness checks. In every case the MV-1/N model is preferable to the other three procedures, which is consistent with our hypothesis.

MV Model*		$\lambda$	$\alpha$	Initial Window	Constraints	Performance Measure	1	2	3	4
							MV-1/N	1/N-1/N	MV-MV	1/N-MV
1	Bayes-Stein	2	0.15	12	VBCs	CER	0.1124	0.1113	0.0913	0.0932
						Sharpe Ratio	0.7974	0.7673	0.6599	0.6815
2	Bayes-Stein	10	0.15	12	VBCs	CER	0.0513	0.0328	0.0436	0.0342
						Sharpe Ratio	0.8375	0.7673	0.7492	0.7108
3	Bayes-Stein	5	0.10	12	VBCs	CER	0.0897	0.0818	0.0787	0.0745
						Sharpe Ratio	0.8413	0.7673	0.7610	0.7198
4	Bayes-Stein	5	0.20	12	VBCs	CER	0.0912	0.0818	0.0710	0.0710
						Sharpe Ratio	0.8487	0.7673	0.6942	0.6936
5	Bayes-Stein	5	-----	12	UGCs	CER	0.0874	0.0818	0.0647	0.0702
						Sharpe Ratio	0.8150	0.7673	0.6425	0.6877
6	Black-Litterman	5	0.15	12	VBCs	CER	0.0873	0.0818	0.0804	0.0743
						Sharpe Ratio	0.8350	0.7673	0.7909	0.7198
7	Markowitz	5	0.15	12	VBCs	CER	0.0838	0.0818	0.0630	0.0659
						Sharpe Ratio	0.7862	0.7673	0.6291	0.6523
8	Bayes-Stein	5	0.15	36	VBCs	CER	0.0745	0.0666	0.0579	0.0583
						Sharpe Ratio	0.7410	0.6793	0.6111	0.6151
9	Bayes-Stein	5	0.15	60	VBCs	CER	0.0620	0.0541	0.0462	0.0458
						Sharpe Ratio	0.6733	0.6118	0.5358	0.5372
10	Bayesian Diffuse Prior	5	0.15	12	VBCs	CER	0.0840	0.0818	0.0634	0.0661
						Sharpe Ratio	0.7878	0.7673	0.6322	0.6537

Table 6: CERs and Sharpe Ratios for the Ten Robustness Checks

---

<sup>11</sup> UGC refers to the homogeneous upper constraints imposed on asset weights as  $x_i \leq (N+1)/2N$  for all  $i$ , where  $N$  is the number of assets. (Platanakis et al., 2019)

\* All the models allow for transactions costs and rule out short sales.

Table 7 presents significance tests on the differences in performance between the four pairs of procedures, again using the tests of Jobson and Korkie (1981), as modified by Memmel (2003), for the Sharpe ratios; and the DeMiguel et al. (2009a) test for the CERs. The results of these tests confirm those in Table 4 and Figure 2, with MV-1/N significantly superior to 1/N-MV and MV-MV; and 1/N-1/N superior to 1/N-MV. In every case in Tables 4 and 7 the statistically significant results are the same for both the Sharpe ratios and CERs.

	Superior Inferior	MV-1/N 1/N-MV	MV-1/N MV-MV	1/N-1/N 1/N-MV	MV-1/N 1/N-1/N	1/N-1/N MV-MV	MV-MV 1/N-MV
1. BS, 2, 0.15, VBCs, 12	SR	0.062*	0.009***	0.047**	0.354	0.173	0.401
	CER <sup>†</sup>	0.043**	0.006***	0.027**	0.401	0.138	0.410
2. BS, 10, 0.15, VBCs, 12	SR	0.071*	0.052*	0.117	0.224	0.439	0.335
	CER <sup>†</sup>	0.072*	0.046**	0.174	0.245	0.445	0.342
3. BS, 5, 0.10, VBCs, 12	SR	0.040**	0.034**	0.114	0.162	0.476	0.283
	CER <sup>†</sup>	0.041**	0.022**	0.107	0.224	0.405	0.322
4. BS, 5, 0.20, VBCs, 12	SR	0.046**	0.008***	0.089*	0.207	0.287	0.493
	CER <sup>†</sup>	0.041**	0.006***	0.086*	0.244	0.266	0.497
5. BS, 5, ---, UGCs, 12	SR	0.127	0.010*	0.092*	0.342	0.202	0.356
	CER <sup>†</sup>	0.115	0.009***	0.088*	0.364	0.196	0.358
6. BL, 5, 0.15, VBCs, 12	SR	0.068*	0.186	0.147	0.214	0.414	0.195
	CER <sup>†</sup>	0.085*	0.119	0.133	0.323	0.461	0.275
7. M, 5, 0.15, VBCs, 12	SR	0.045**	0.006***	0.016**	0.407	0.116	0.394
	CER <sup>†</sup>	0.042**	0.005***	0.019**	0.428	0.114	0.395
8. BS, 5, 0.15, VBCs, 36	SR	0.078*	0.014**	0.106	0.261	0.289	0.485
	CER <sup>†</sup>	0.077*	0.013**	0.147	0.280	0.306	0.488
9. BS, 5, 0.15, VBCs, 60	SR	0.086*	0.019**	0.086*	0.280	0.277	0.493
	CER <sup>†</sup>	0.088*	0.019**	0.157	0.294	0.327	0.486
10. BDP, 5, 0.15, VBCs, 12	SR	0.045**	0.006***	0.017**	0.400	0.121	0.402
	CER <sup>†</sup>	0.041**	0.005***	0.02**	0.421	0.118	0.402

Table 7: Probability Values for the Significance Tests of the Differences Between Sharpe Ratios and CERs for the Four Two-stage Procedure Robustness Checks

Significance at the 10%, 5% and 1% levels is denoted by \*, \*\* and \*\*\* respectively

† The significance tests for CERs use the value of  $\lambda$  used in forming the portfolio

As a further robustness check, we repeated our the two stage methodology of our core analysis on international data for the UK, USA, Germany, Switzerland, France, Canada and Brazil. This consists of value-weighted total return equity market indices for seven countries - UK (FTSE 100), US (S&P500), Germany (DAX 30), Switzerland (SMI), France (CAC 40), Canada (S&P/TSX Composite), and Brazil (Bovespa); with monthly data from December 1994 to August 2017 expressed in \$US. We also analysed the ten companies with the largest market capitalization in each index in December 1994; so in total we have 70 companies (see Appendix 3). We used 1-month T-bill returns from Ken French's web site as the riskless rate.

Table 8 shows that for asset allocation across these counties MV remains dominant, while Table 9 indicates that  $1/N$  is generally preferred for stock selection, although not for Switzerland or Brazil. Table 10 confirms our main hypothesis, as MV- $1/N$  is the best strategy for selecting international portfolios. We also find that, on balance, its reverse,  $1/N$ -MV, is the worst strategy.

Performance Measure	1 $1/N$	2 MV
CER	0.0126	0.0272
Sharpe Ratio	0.4178	0.4233
Dowd Ratio	0.2328	0.2398
Omega Ratio	1.4994	1.5305

Table 8: Performance of  $1/N$  and MV in Forming Portfolios of the Seven Countries, i.e. Stage One Asset Allocation - 12-month expanding estimation window, December 1994 to August 2017,  $\lambda = 5$ , Transaction costs = 50 bps.

Countries	CER		Sharpe Ratio		Dowd Ratio		Omega Ratio	
	$1/N$	MV	$1/N$	MV	$1/N$	MV	$1/N$	MV
UK	0.0446	0.0345	0.5480	0.4632	0.3326	0.2690	1.6889	1.5755
US	0.0731	0.0615	0.7182	0.6275	0.4989	0.4117	1.9767	1.8392
Germany	-0.0444	-0.0495	0.3660	0.2818	0.1951	0.1442	1.4163	1.3261
France	0.0062	0.0144	0.4599	0.4021	0.2603	0.2229	1.5338	1.4755
Canada	0.0619	0.0535	0.6736	0.6181	0.4388	0.3897	1.8320	1.7582
Switzerland	0.0113	0.0432	0.4609	0.5274	0.2617	0.3173	1.5487	1.6300
Brazil	-0.2499	-0.1657	0.5368	0.5759	0.3077	0.3386	1.5987	1.6363

Table 9: Performance of  $1/N$  and MV in Forming Portfolios of the Shares Within Each County, i.e. Stage Two Stock Selection - 12-month expanding estimation window, December 1994 to August 2017,  $\lambda = 5$ , Transaction costs = 50 bps.

Performance Measure	1 MV-1/N	2 1/N-1/N	3 MV-MV	4 1/N-MV
CER	0.0632	0.0510	0.0590	0.0552
Sharpe Ratio	0.6512	0.6232	0.6110	0.6117
Dowd Ratio	0.4244	0.3922	0.3928	0.3863
Omega Ratio	1.8318	1.7642	1.7783	1.7554

Table 10: Performance Measures for the Four Two-stage Procedures for the International Data - 12-month Expanding Estimation Window, December 1994 to August 2017,  $\lambda = 5$ , Transaction costs = 50 bps.

## 8. Conclusions

For a range of organizational reasons, large investors typically split their portfolio decision into two stages - asset allocation and stock selection. We hypothesise that mean-variance models are superior to  $1/N$  for asset allocation, while the reverse applies for stock selection. This is because estimation errors are lower for asset classes than for individual assets, and so are less of a problem for mean-variance models when used for asset allocation than stock selection. We identify three reasons for the superiority of mean-variance over  $1/N$  for asset allocation. First, a shorter time series is required to estimate the covariance matrix, particularly for asset classes, which offers more degrees of freedom when estimating the inputs for asset allocation. Second, asset classes are grouped data, resulting in smaller estimation errors for asset classes than individual assets. Finally, there are fewer asset classes than individual assets, leading to a smaller chance of outliers and the error maximization problem for asset allocation.  $1/N$  is unaffected by estimation errors, and so its performance, relative to mean-variance, is greater for stock selection than asset allocation.

Using a sample of US equities for ten asset classes and the 93 underlying individual assets, we confirm this hypothesis using four different types of mean-variance model and a range of parameter settings; suggesting that our hypothesis applies across a wide range of mean-variance models. We also replicate the simulation analysis of DeMiguel et al. (2009a) for US data which compared the performance of mean-variance and  $1/N$ , but relax their assumption that equities have  $\alpha$ s that are strictly zero. When we do this we find the conclusions of DeMiguel et al. (2009a) are reversed - the smaller is idiosyncratic



asset volatility, the better mean-variance performs relative to  $1/N$ . This finding is consistent with our empirical results that mean-variance is superior to  $1/N$  for asset classes. Finally we calibrate the simulation analysis to our sample of US equities, and show that the Markowitz and Bayes-Stein models are superior to  $1/N$  for asset classes, while  $1/N$  is superior for stock selection. We support our conclusions with ten robustness checks with the US data, and the replication of our core analysis on international data from seven countries.

## References

- Allaj, E. (2017) The Black-Litterman Model and Views from a Reverse Optimization Procedure: An Out-of-Sample Performance Evaluation, Working paper, University of Florence.
- Barroso, P. (2015) Using Out-of-Sample Errors in Portfolio Optimization, Working paper, University of New South Wales.
- Barry, C. B. (1974) Portfolio Analysis under Uncertain Means, Variances, and Covariances, *Journal of Finance*, vol. 29, pp. 515–522.
- Bessler, W., Opfer H. and Wolff, D. (2017) Multi-Asset Portfolio Optimization and Out-of-Sample Performance: An Evaluation of Black-Litterman, Mean Variance and Naïve Diversification Approaches, *European Journal of Finance*, vol. 23, no. 1, pp. 1-30.
- Bessler, W. and Wolff, D. (2015) Do Commodities Add Value in Multi-Asset Portfolios? An Out-of-Sample Analysis for Different Investment Strategies, *Journal of Banking and Finance*, vol. 60, pp. 1-20.
- Black F. and Litterman R. (1992) Global Portfolio Optimization, *Financial Analysts Journal*, vol. 48, no. 5, pp. 28–43.
- Blake, D., Rossi, A.G., Timmermann, A., Tonks, I. and Wermers, R. (2013) Decentralized Investment Management: Evidence from the Pension Fund Industry, *Journal of Finance*, vol. 68, no. 3, June, pp. 1133-1178.
- Board, J.L.G. and Sutcliffe, C.M.S. (1994) Estimation Methods in Portfolio Selection and the Effectiveness of Short Sales Restrictions: UK Evidence, *Management Science*, vol. 40, no. 4, April, pp. 516-534.
- Brinson, G.P., Hood, L.R. and Beebower, G.L. (1986) Determinants of Portfolio Performance, *Financial Analysts Journal*, vol. 42, no. 4, July-August, pp. 39-44.
- DeMiguel, V., Garlappi, L. and Uppal, R. (2009a) Optimal Versus Naive Diversification: How Inefficient is the  $1/N$  Portfolio Strategy?, *Review of Financial Studies*, vol. 22, no. 5, May, pp. 1915-1953.
- DeMiguel, V., Garlappi, L., Nogales, F.J. and Uppal, R. (2009b) A Generalized Approach to Portfolio Optimization: Improving Performance by Constraining Portfolios, *Management Science*, vol. 55, no. 5, May, pp. 798-812.
- DeMiguel, V., Martin Utrera, A. and Nogales, F.J. (2013) Size Matters: Optimal Calibration of Shrinkage Estimators for Portfolio Selection, *Journal of Banking and Finance*, vol. 37, no. 8, pp. 3018-3034.
- Dickson, M. (2016) Naive Diversification isn't so Naive After All, Working paper, University of North Carolina.
- Diris, B., Palm, F. and Schotman, P. (2015) Long-Term Strategic Asset Allocation: An Out-of-Sample Evaluation, *Management Science*, vol. 61, no. 9, pp. 2185-2202.
- Durand, R.B., Gould, J. and Maller, R. (2011) On the Performance of the Minimum Portfolio, *European Journal of Finance*, vol. 17, no. 7, August, pp. 553-576.
- Elton, E.J. and Gruber, M.J. (1971) Improved Forecasting Through the Design of Homogeneous Groups, *Journal of Business*, vol. 44, no. 4, October, pp. 432-450.
- Frankfurter, G.M. and Phillips, H.E. (1980) Portfolio Selection: An Analytic Approach for Selecting Securities from a Large Universe, *Journal of Financial and Quantitative Analysis*, vol. 15, no. 2, June, pp. 357-377.
- Han, C. (2016) Improving the Naive Portfolio Strategy, Working paper, University of Durham.
- Hanoch, G. and Levy, H. (1969) The Efficiency Analysis of Choices Involving Risk, *Review of Economic Studies*, vol. 36, no. 3, pp. 335-346.
- Hwang, I., Xu, S. and In, F. (2018) Naive Versus Optimal Diversification: Tail Risk and Performance, *European Journal of Operational Research*, vol. 265, no. 1, February, pp. 372-388.
- Ibbotson, R.G. and Kaplan, P.D. (2000) Does Asset Allocation Policy Explain 40, 90, or 100 Percent of Performance?, *Financial Analysts Journal*, vol. 56, no. 1, January-February, pp. 26-33.
- Jagannathan, R. and Ma, T. (2003) Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps, *Journal of Finance*, vol. 58, no. 4, August, pp. 1651-1683.
- James, W., and C. Stein. (1961) Estimation with Quadratic Loss, *Proceedings of the 4th Berkeley Symposium on Probability and Statistics*, vol. 1, pp. 361-379.
- Jarrow, R. (2010) Active Portfolio Management and Positive Alphas: Fact or Fantasy?, *Journal of Portfolio Management*, vol. 36, no. 4, Summer, pp. 17-22.
- Jarrow, R. and Zhao, F. (2006) Downside Loss Aversion and Portfolio Management, *Management Science*, vol. 52, no. 4, April, pp. 558-566.
- Jobson, J.D., & Korkie, B. (1980). Estimation for Markowitz Efficient Portfolios. *Journal of the American Statistical Association*, vol. 75, no. 317, September, pp. 544-554.
- Jobson, J.D. and Korkie, B.M. (1981) Performance Hypothesis Testing with the Sharpe and Treynor Measures, *Journal of Finance*, vol. 36, no. 4, September, pp. 889-908.
- Jorion, P. (1985) International Portfolio Diversification with Estimation Risk, *Journal of Business*, vol. 58, no. 3, pp. 259-78.

- Jorion, P. (1986) Bayes-Stein Estimation for Portfolio Analysis, *Journal of Financial and Quantitative Analysis*, vol. 21, no. 3, September, pp. 279-292.
- Kan, R. and Zhou, G. (2007) Optimal Portfolio Choice with Parameter Uncertainty, *Journal of Financial and Quantitative Analysis*, vol. 42, no. 3, pp. 621-656.
- Kan, R., Wang, X. and Zhou, G. (2016) Optimal Portfolio Choice with Estimation Risk: The Case without a Risk-free Asset, *Rotman School of Management Working Paper No. 2819254*.
- Kirby, C. and Ostdiek, B. (2012) It's All in the Timing: Simple Active Portfolio Strategies that Outperform Naïve Diversification, *Journal of Financial and Quantitative Analysis*, vol. 47, No. 2, pp. 437-467.
- Klein, R. W. and Bawa, V. S. (1976) The Effect of Estimation Risk on Optimal Portfolio Choice, *Journal of Financial Economics*, vol. 3, pp. 215-31.
- Kolm, P.N., Tutuncu, R. and Fabozzi, F.J. (2014) 60 Years of Portfolio Optimization: Practical Challenges and Current Trends, *European Journal of Operational Research*, vol. 234, no. 2, pp. 356-371.
- Lam, K., Mok, H.M.K., Cheung, I. and Yam, H.C. (1994) Family Groupings on Performance of Portfolio Selection in the Hong Kong Stock Market, *Journal of Banking and Finance*, vol. 18, no. 4, pp. 725-742.
- Li, B. (2016) Sample Selection Bias Return Moments and the Performance of Optimal Versus Naive Diversification, Working paper, University of Melbourne.
- Markowitz, H. (1952) Portfolio Selection, *Journal of Finance*, vol. 7, no. 1, March, pp. 77-91.
- Mommel, C. (2003) Performance Hypothesis Testing with the Sharpe Ratio, *Finance Letters*, vol. 1, No. 1, pp. 21-23.
- Meucci, A. (2010) The Black-Litterman Approach: Original Model and Extensions, *Encyclopedia of Quantitative Finance*, Wiley.
- Michaud, R.O. and Michaud, R.O. (2008) *Efficient Asset Management: A Practical Guide to Stock Portfolio Optimization*, Oxford University Press, Second edition.
- Levy, H. and Levy, M. (2014) The Benefits of Differential Variance-Based Constraints in Portfolio Optimization, *European Journal of Operational Research*, vol. 234, no. 2, April, pp. 372-381.
- Levy, M. and Roll, R. (2010) The Market Portfolio May Be Mean-Variance Efficient After All, *Review of Financial Studies*, vol. 23, no. 6, pp. 2464-2491.
- Oikonomou, I., Platanakis, E. and Sutcliffe, C. (2018) Socially Responsible Investment Portfolios: Does the Optimization Process Matter? *British Accounting Review*, vol. 50, no. 4, pp. 379-401.
- Platanakis, E., Sakkas, A., Sutcliffe, C. (2019) Harmful Diversification: Evidence from Alternative Investments, *British Accounting Review*, vol. 51, no. 1, January, pp. 1-23.
- Platanakis, E. and Sutcliffe, C. (2017) Asset-Liability Modelling and Pension Schemes: The Application of Robust Optimization to USS, *European Journal of Finance*, vol. 23, no. 4, pp. 324-352.
- Platanakis, E. and Urquhart, A. (2019) Portfolio Management with Cryptocurrencies: The Role of Estimation Risk, *Economics Letters*, vol. 177, April, pp. 76-80.
- Prigent, J.L. (2007) *Portfolio Optimization and Performance Analysis*, Chapman and Hall.
- Satchell, S. and Scowcroft, A. (2000) A Demystification of the Black-Litterman Model: Managing Quantitative and Traditional Portfolio Construction, *Journal of Asset Management*, vol. 1, no. 2, pp. 138-150.
- Shadwick, W.F. and Keating, C. (2002). A Universal Performance Measure. *Journal of Performance Measurement*, vol. 6, pp. 59-84.
- Sharpe, W. F. (1966). Mutual Fund Performance. *Journal of Business*, 39, 119-138.
- Sharpe, W.F. (1981) Decentralized Investment Management, *Journal of Finance*, vol. 36, no. 2, May, pp. 217-234.
- Silva, T., Pinheiro, P.R. and Poggi, M. (2017) A More Human-Like Portfolio Optimization Approach, *European Journal of Operational Research*, vol. 256, no. 1, pp. 252-260.
- Stracca, L. (2006) Delegated Portfolio Management: A Survey of the Theoretical Literature, *Journal of Economics Surveys*, vol. 20, 5, December, pp. 823-848.
- Shigeta, Y. (2016) Optimality of Naive Investment Strategies in Dynamic Mean-Variance Optimization Problems with Multiple Priors, Working paper, Kyoto University.
- Tu, J. and Zhou, G. (2011) Markowitz Meets Talmud: A Combination of Sophisticated and Naive Diversification Strategies, *Journal of Financial Economics*, vol. 99, no. 1, pp. 204-215.
- Van Binsbergen, J.H., Brandt, M.W. and Koijen, R.S.J. (2008) Optimal Decentralized Investment Management, *Journal of Finance*, vol. 63, no. 4, August, pp. 1849-1895.
- Welch, I. and Goyal, A. (2008) A Comprehensive Look at The Empirical Performance of Equity Premium Prediction, *Review of Financial Studies*, vol. 21, no. 4, pp. 1455-1508.

## Appendix 1: S&P Sectors and the Largest Constituents

Sector	Constituents*
SP4EFIN Financials	JPM, BAC, WFC, C (Citi Group), MS, AXP(American Express), PNC, AIG, MMC(Marsh & McLennan Cos.), ALL(Allstate Corp.)
SP5ETEL Telecommunications	T(AT&T Inc.), VZ(Verizon Communications Inc.), CTL(CenturyLink Inc.)
SP5EHCR Health Care	JNJ(Johnson & Johnson), PFE(Pfizer Inc.), UNH(UnitedHealth Group Inc.), MRK(Merck & Co. Inc.), AMGN(Amgen Inc.), BMY(Bristol-Myers Squibb Co.), ABT(Abbott Laboratories), LLY(Eli Lilly & Co.), AET(Aetna Inc.), BAX(Baxter International Inc.)
SP5ECST Consumer Staples	WMT, PG(Procter & Gamble Co.), KO(Coca-Cola Co.), PEP(PepsiCo Inc.), MO(Altria Group Inc.), CL(Colgate-Palmolive Co.), KMB(Kimberly-Clark Corp.), GIS(General Mills Inc.), K(Kellogg Co.), CPB(Campbell Soup Co.)
SP5EIND Industrials	GE, BA(Boeing Co.), MMM(3M), HON(Honeywell International Inc.), UNP(Union Pacific Corp.), CAT(Caterpillar Inc.), RTN(Raytheon Co.), CSX(CSX Group), EMR(Emerson Electric Co.), NSC(Norfolk Southern Corp.)
SP5EMAT Materials	PX(Praxair Inc.), SHW(Sherwin-Williams Co.), APD(Air Products & Chemicals Inc.), PPG(PPG Industries Inc.), IP(International Paper Co.), NEM(Newmont Mining Corp.), NUE(Nucor Corp.), FMC(FMC Corp.), EMN(Eastman Chemical Co.), IFF(International Flavors & Fragrances Inc.)
SP5ECOD Discretionary Consumption	HD(Home Depot Inc.), DIS(Walt Disney Co.), MCD(McDonald's Corp.), F(Ford Motor Co.), CCL(Carnival Corp.), TGT(Target Corp.), GPC(Genuine Parts Co.), LB(L Brands Inc.), GPS(Gap Inc.), GT(Goodyear Tire & Rubber Co.)
SP5EUTL Utilities	NEE(NextEra Energy Inc.), DUK(Duke Energy Corp.), D(Dominion Energy Inc.), SO(Southern Co.), AEP(American Electric Power Co. Inc.), EXC(Exelon Corp.), PCG(PG&E Corp.), ED(Consolidated Edison Inc.), EIX(Edison International), PEG(Public Service Enterprise Group Inc.)
SP5EENE Energy	XOM(Exxon Mobil Corp.), CVX(Chevron Corp.), SLB(Schlumberger Ltd.), COP(ConocoPhillips), EOG(EOG Resources Inc.), OXY(Occidental Petroleum Corp.), HAL(Halliburton Co.), BHGE(Baker Hughes, a GE co.), HES(Hess Corp.), MRO(Marathon Oil Corp.)
SP5EINT Information Technology	MSFT, ORCL(Oracle Corp.), INTC(Intel Corp.), CSCO(Cisco Systems Inc.), IBM, ADP(Automatic Data Processing Inc.), HPQ(HP Inc.), CA(CA Inc.), MSI(Motorola Solutions Inc.), XRX(Xerox Corp.)

\* All returns are from Datastream.

## Appendix 2: Performance Measures

The Sharpe ratio (Sharpe, 1966) is one of the most popular metrics for measuring risk-adjusted returns, and is defined as:-

$$SR = (\bar{\mu}_p - \bar{r}_f) / \sigma_p$$

where  $(\bar{\mu}_p - \bar{r}_f)$  represents the out-of-sample mean excess portfolio return, and  $\sigma_p$  is the portfolio standard deviation over the entire out-of-sample (investment) period. The Sharpe ratio has its limitations, and for this reason we also employ several other metrics.

The Certainty Equivalent Return (CERs) for mean-variance investors can be approximated and computed as follows:-

$$CER = \bar{\mu}_p - (\lambda / 2) \sigma_p^2$$

where  $\lambda$  is the relative risk aversion parameter, and  $\bar{\mu}_p$  and  $\sigma_p^2$  have been defined above.

The Omega ratio (Shadwick and Keating, 2002) with a target return of zero, also known as the average gain to the average loss ratio, is computed as follows:-

$$Omega = [(1/T) \sum_{t=1}^T \max(0, +R_{p,t})] / [(1/T) \sum_{t=1}^T \max(0, -R_{p,t})]$$

where  $R_{p,t}$  denotes the out-of-sample portfolio return at time  $t$ . The main advantage of the Omega ratio is that it does not require any assumption about the distribution of portfolio returns.

The Dowd ratio is the out-of-sample mean excess portfolio return, divided by the portfolio value-at-risk (Prigent, 2007), and is computed as follows:-

$$Dowd = (\bar{\mu}_p - \bar{r}_f) / VaR_{95\%}$$

The VaR in the Dowd ratio has been computed at the 95% confidence level over the entire investment (out-of-sample) period.

### Appendix 3: Largest Ten Companies in the Seven Country Indices

Country	Constituents
UK - FTSE100	Barclays Bank, British Petroleum, Unilever, GlaxoSmithKlein, British American Tobacco, Diageo, Rio Tinto, Royal Dutch Shell, British Telecommunications, Marks & Spencer
USA - S&P500	Walmart, Exxon-Mobil, Coca-Cola, IBM, General Electric, Proctor & Gamble, Merck, Pepsico, Altria, Bristol Myers Squibb.
Germany - DAX 30	Deutsche Bank, BMW, Allianz, Siemens, BASF, Bayer, RWE, Munich Re, E.ON, ThyssenKrupp.
France - CAC 40	BNP Paribas, L'Oreal, Total, Societe Generale, AXA, Danone, LVMH, Air Liquide, Carrefour, Vivendi.
Canada - S&P/TSX Composite	Royal Bank of Canada, Bank of Montreal, Toronto Dominion Bank, Bank of Nova Scotia, Canadian Imperial Bank of Commerce, Bell Canada Enterprises, Imperial Oil, Barrick Gold, Encana, Thomson-Reuters.
Switzerland - SMI	Nestle, UBS, Roche, Credit Suisse, Novartis, ABB, Zurich Insurance, Richemont, Swiss Re, Swatch.
Brazil - Bovespa	Vale, Petrobras, Companhia Siderúrgica Nacional, Usiminas, Eletrobras, CEMIG, ITAU Unibanco, Banco do Brazil, Bradesco, Lojas Americanas